

1. Introduction

The Global River Radar Altimeter Time Series (GRRATS) are river heights from OSTM/Jason-2 and Envisat that are conformed to look like river gauges via virtual stations (VS). The purpose of these heights are to provide satellite altimetric river height data in a form that is more recognizable to the observational community and as a way to get users use to using satellite data for river hydrology. GRRATS provides data from 932 VS on 39 of the world's largest rivers (wider than 900m). River heights were processed with limits established by DEM data from inside the VS. When applicable, times of ice cover are also flagged consistently. To allow for maximum usability, all processing data are included (original L2 data, filtering limits, etc.). When possible, data were validated with in situ gauges. Other locations were assigned a qualitative letter grade, based on the amount of missing data, agreement with nearby VS and identifiable seasonal cycle. Validation information (quantitative or qualitative) is packaged with each VS's data to aid the end user in selection the best time series for their particular task.

Motivation for GRRATS

There are currently four widely-used radar altimeter data products for inland areas. The European Space Agency (ESA) Global Near-Real Time River and Lake product (NRTRL hereafter) (P. Berry & Wheeler, 2009), the Database for Hydrological Time Series of Inland Waters (DAHITI) (Schwatke et al., 2015), the Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS) Hydroweb product (Santos da Silva et al., 2010), and the University of Stuttgart, Institute of Geodesy (GIS) product, HydroSat (Tourian et al., 2016). Radar altimeter processing for rivers involves seven steps: site selection, extraction of the radar returns over the river, retracking waveforms, applying corrections, post facto filtering, QA/QC, and data distribution. Methods used to construct each data product and details about them are shown in Table 1 and Table 2; note that many of the methods are not described.

Many existing products are not well documented. From Table 1, only NRTRL provided any description of their site selection methodology, VS were included in the database if they “produc[ed] good results” (P. Berry & Wheeler, 2009). The total number of VS on rivers varies significantly and some products do not break out water level VS by water body type (i.e. rivers vs lakes), so the number of river VS for those products is unknown. From Table 2, only Hydroweb describes both the radar waveform retracking methods and geophysical corrections used in creating the dataset. Return selection and filtering methods sometimes go undescribed and it is not always clear how uniformly these methods are applied across the dataset.

Table 1 Product VS totals and user selection criteria: Column 2 indicates if the VS are for lakes (L) or rivers(R). Column 3 indicates if the VS are grouped separately in the user interface. Column 4 indicates if the entire product is available as a bulk download

Product	Potential temporal range	Total VS (from rivers)	Rivers and lakes separate	Bulk Download	User Interface	Site Selection
ESA NRT Rivers and lakes	2002-2016	1910 (N/A)	no	no	Map	Good results
DAHITI	1992-present	440 (293)	yes*	no	Map/ browse	-
Hydroweb	2002-present	1406 (1259)	yes	yes	Map/ browse/ select by	-
HydroSat	1992-present	683 (N/A)	yes*	no	Map	-
GRRATS	2002-2016	920 (920)	N/A	yes	Map/FTP	all ocean-draining rivers width > 900 m

No existing products have uniform validation of their VS published in peer-reviewed literature. Moreover, none of the existing products contains any kind of validation information with their time series. HydroSat includes the formal statistical uncertainty of each point resulting from its densification algorithm, but not validation information. With the exception of NRTRL, all of the datasets are validated in published case studies, but the extent to which the remainder of the dataset underwent validation is unclear. No validation information is included with the data the end user accesses. This makes the data use challenging, as the quality of each VS is unknown.

Table 2 radar altimeter product processing (* indicates the item is not described.)

Product	Retracking/ corrections described	Return selection	Filtering	Ice	QA/QC	Traceability/ Transparency
NRTRL	*	Water Mask Backscatter	*	*	*	Final Product only
Dahiti	Corrections only	*	Backscatter / statistical methods	*	Case study	Final Product only
Hydroweb	both	Polyline segment/google earth overlay	Visual inspection/ physically based corrections/ statistical methods	*	Case study	Final Product only
Hydrosat	*	Narrow river water mask (Landsat/RivWidth)	Statistical methods	*	Statistical error for each point	Final Product only
GRRATS	both	Median flow water mask (Landsat from GRWL) with 2km buffer	DEM limit/ statistical methods	A priori and remote sense data	Included for each VS	All production components included

None of the existing datasets have global coverage (Table 3). The existing products do not include data from Arctic rivers and are relatively sparse in their coverage of North America. It is likely that Arctic rivers are excluded due to issues related to ice cover and lack of in-situ data for quantitative validation, although (as noted in Table 1), the rationale for VS inclusion in existing datasets is not described. It is also important to note that altimeter coverage of high latitudes is limited (platforms on the 10 day TOPEX/Poseidon orbit only extend to 66° latitude).

Many of the existing products are not easily accessible for the end user. From Table 1, with the exception of Hydroweb, none of the products offers bulk download, and none offers a database of VS locations, for easily querying the dataset. The existing products primarily offer a map interface where the user clicks the location they are interested in to download a time series. In many cases the number and nature of the VS (river or lake) is obscured by several clicks, and in the case of NRTRL, rivers and lakes are not differentiated in the metadata. These issues substantially affect usability for researchers trying to ingest large amounts of data.

Table 3 Number of virtual stations in GRRATS after filtering. The percentage of possible VS (intersections of Jason-2 and Envisat ground tracks with 900 m or wider sections of the river) that remained in the product after processing is indicated in column 4. Column five indicates if a river present in GRRATS is missing in existing datasets. DH,HW,and HS were used to abbreviate DAHITI, Hydroweb, and HydroSat respectively. NRTRL was not available at the time of this product comparison. Column 6 indicates the presence of a location of overlap where direct comparison of all of the products was possible.

River	Jason-2	Envisat	% of possible VS in final product	Not present in	Overlap location
North America					
Columbia	3	3	85.71	DH, HW	-
Mackenzie	9	24	34.37	DH, HW	-
Mississippi	5	11	69.56	DH, HW	-
St Lawrence	2	11	81.25	DH, HW,HS	-
Susquehanna	2	1	60	DH, HW,HS	-
Yukon	19	34	72.6	HW	-
South America					
Amazon	21	80	86	*	Yes
Courantyne	0	1	100	DH, HW,HS	-
Essequibo	0	1	100	DH, HW,HS	-
Magdalena	1	0	33.33	HS	-
Oiapoque	1	1	100	HW,HS	-
Orinoco	8	29	86.05	*	No

Parana	21	28	90.74	HW	-
Sao Francisco	1	4	71.43	HW	-
Tocantins	2	8	52.63	HW	-
Uruguay	4	10	93.33	DH, HW,HS	-
Africa					
Congo	12	63	84.26	*	Yes
Niger	1	5	66.67	*	No
Nile	1	2	25	HW,HS	-
Zambezi	4	5	64.29	DH, HW,HS	-
Europe					
Mezen	1	4	100	DH,HW	-
Pechora	0	6	11.11	DH,HW	-
Volga	12	24	70.59	*	No
Asia					
Amur	10	47	96.66	DH,HW	-
Anabar	0	5	71.43	DH,HW	-
Anadyr	8	10	78.26	DH,HW	-
Ayeyarwada	3	10	92.85	*	No
Brahmaputra	3	16	82.61	*	Yes
Ganges	6	16	88	*	No
Indus	1	12	86.67	*	No
Khatanga	0	15	78.95	DH,HW,HS	-
Kolyma	0	28	58.33	DH,HW	-
Kuloy	0	4	100	DH,HW,HS	-
Lena	2	21	10.84	DH	-
Mekong	0	11	64.71	*	No
Ob	34	53	73.11	DH,HW	-
Olenyok	0	1	9.09	DH,HW,HS	-

Pyasina	0	6	46.15	DH,HW,HS	-
Yangtze	7	28	83.33	DH,HW	-
Yenisei	26	54	66.66	DH,HW	-

Pulse limited radar altimeter footprint size

In previous sections, citations were given to demonstrate how large the Envisat and Jason-2 footprints could be. However, the footprint of radar altimeter as described in terms of a pulse-limited footprint varies in size primarily due to significant wave height on the measured surface. From Fu and Cazenave (2000) we present equation (1) for the radius of an altimeter footprint is:

$$r(H_w) = \left[\frac{(c\tau + 2H_w)R_0}{1 + R_0/R_e} \right]^{1/2}$$

(1)

Where, c is the speed of light R_e is the radius of the earth, ($R_e = 6371 \text{ km}$), $R_0 = c * t_0$, t_0 is the arrival time of the first radar return from the wave crest at nadir $c\tau$ is the width of the pulse, the effective pulse duration that has been used in Jason-2 and Envisat altimeters is $\tau = 3.125 \text{ nsec}$ (S. Gleason et al., 2010). H_w is the crest-to-trough wave height for a special consideration that sea surface consisting of a monochromatic, unidirectional wave-train. The H_w can be replaced by the significant wave height, $H_{1/3}$. $H_{1/3}$ is usually considered to be equivalent to four times the standard deviation of the wave height distribution. With a significant wave height approaching zero, as would be expected in fluvial systems, it is likely that footprint sizes will be much smaller, around 1.6-1.7km (Fetterer et al., 1992; Sulistioadi et al., 2015). Any calculated footprint size on a river will always be an estimation. Equation (1) is really only applicable for calculating a footprint size over the ocean.

A note about temporal resolution

All of the altimeter missions used in the construction of GRRATS follow one of two orbits. This means that they will have an approximately 10 day resolution (TOPEX/Poseidon, Jason-1, Jason-2, Jason-3), or and approximately 35 day resolution (ERS-1, ERS-2, Envisat).

2. Datasets used

We compiled GRRATS using Envisat and Jason-2 radar altimetry data (sections 2.1 and 2.2). We later added ERS-1, ERS-2, TOPEX/Poseidon, Jason-1 and Jason-3 in GRRATS version 2 (sections 2.3.1-2.3.5). We used the Global River Widths from Landsat (GRWL) database to identify the locations where the altimeters crossed ocean draining rivers with widths >900m (section 2.3). In order to flag altimeter returns from ice-covered rivers we used river ice data acquired from United States Geological Survey (USGS) and Environment and Climate Change Canada (ECCC) river gage system, as well as existing

literature (section 2.4). We used Shuttle Radar Topography Mission (SRTM), Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010), and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Digital Elevation Model (DEM) data (sections 2.5, 2.6, and 2.7) to establish approximate expected river elevations for use in filtering. We used stage from primarily public sources, including USGS, ECCC, the Mekong river commission (MRC), Geodynamical, hydrological and biogeochemical control of erosion/alteration and material transport in the Amazon, Orinoco and Congo basins (SO HYBAM) and private sources for validation (described in section 3.8).

2.1 Envisat

The European Space Agency (ESA) Envisat mission began in February 2002 and ended in April 2012. Envisat followed a sun-synchronous, 35-day repeat orbit, with a 98.4° inclination. Envisat used the Radar Altimeter 2 (RA-2) instrument to measure the range between Envisat and the target surface (through time delay) and the target surface's characteristics (through the power and shape of the return). The RA-2 footprint was ~3.4 km in diameter (Frappart et al., 2006). It operated with 13.575 GHz as its primary frequency and 3.2 GHz to allow for corrections (Louet & Bruzzi, 1999). Envisat data are available at both 1 and 18 Hz frequency, where 1 or 18 readings are included per flight second. Level-2 Envisat data products provide geo-located values for radiance, reflectance, transmittance, polarisation, radar backscattering values, and radar echo-time delay (Louet & Bruzzi, 1999). Editing criteria for GDR data are can be found in Table 4 (Soussi & Féménias, 2009).

2.2 OSTM/ Jason-2

Launched in June of 2008, the OSTM /Jason-2 satellite altimeter was designed to establish a precise multidecadal record of ocean topography. The mission is a joint initiative between the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), European Organization of the Exploitation of Meteorological Satellites (EUMETSAT), and the French space agency, the Centre National d'Études Spatiales (CNES) (Lambin et al., 2010). Jason-2 follows a repeat track on a non-sun-synchronous orbit, with a 66° inclination (meaning that 66° N is the highest latitude the orbit crosses). Cycle length is 9.9 sidereal days. Its orbit could drift up to ±1km from the nominal (repeat) track at the equator (Flohrer et al., 2011).

The on-board altimeter, a dual frequency (Ku and C-band operating at 13.5-GHz and 5.3-GHz, respectively (Desjonquères et al., 2010)) radar altimeter measures the altimetric range, or distance between the instrument and surface at nadir and provides propagation delay correction information. The altimeter footprint is 3 to 4-km in diameter (Papa et al., 2012). The OSTM/Jason-2 level-2 products include geo-located values for the altimetric range, *i.e.* distance between the satellite (radar altimeter phase center) and the water surface, orbit estimation, significant wave height, and backscatter coefficient (Lambin et al., 2010). These data are available at both 1 and 20-Hz. Editing criteria for GDR data are can be found in Table 4 (Dumont et al., 2009).

2.3 Additional Altimeters included in GRRATSV2

2.3.1 ERS-1

Launched by the ESA in July of 1991 and ending in March of 2000 the European Remote-Sensing Satellite-1 (ERS-1) mission was created to assist in a wide range of scientific disciplines including ocean and land cover and meteorological studies. Its orbit was sun synchronous, near-circular, at a 98.52°

inclination and an altitude of 782-785 km. The mapping mode had a 35day repeat orbit. The onboard altimeter was the radar altimeter 1 (RA-1), a Ku-band altimeter which operated at 13.8 GHz and had a theoretical footprint size of ~16-20km. Data are available at both 1 and 20 Hz frequency, where 1 or 18 readings are included per flight second. (Duchossois, 1991; Kramer, 2002; Proud & Battrick, 1992).

2.3.2 ERS-2

The European Remote-Sensing Satellite-2 (ERS-2) mission was a follow up mission to ERS-1. The mission was launched in April of 1995 and ended in September 2011. With the exception of a new instrument designed to measure the chemical composition of the atmosphere, the mission parameters are nearly identical("ERS-2 Product Handbook," 2016; Francis et al., 1995; Kramer, 2002).

2.3.3 TOPEX/POSEIDON

Launched in August 1992 and ending January 2006, The TOPEX/POSEIDON mission was a Joint NASA and CNES initiative. Its primary mission objectives were collecting a three year view of the Earth's oceans, improving understanding of ocean currents and improving understanding of global climate. It flew a circular, non-sun-synchronous orbit with a 66° inclination at an altitude of 1336km and a ~10 day repeat cycle. The primary mission payload was the NASA Radar Altimeter (NRA). It was a dual band 13.6 GHz(KU-band) and 5.3GHz(C-band) altimeter which where both frequencies could be operated simultaneously allowing for estimation of ionosphere delay. The platform also carries the TOPEX Microwave Radiometer (TMR) for accounting for water vapor delay, an experimental GPS receiver (GPS Demonstration Receiver), and experimental solid-state altimeter (the Poseidon-1), and a Laser Reflector Array (LRA) use for ground tracking. Mission data include sea surface topography, significant wave height, surface wind speed, ocean tides, vertically integrated atmospheric water vapor and vertically integrated ionosphere electron content. They are available at 1 and 20hz (Fu et al., 1994; Hupp & Moreaux, 2006; Kramer, 2002). The altimetric data of the TOPEX mission used in GRRATS was extracted from Geophysical Data Record (GDR) in netCDF format, provided by the CASH consortium (IRD, Legos, CLS and BRLi), reprocessed and distributed by Legos/CTOH, 2013(Benada, 1997; CASH consortium (IRD, Legos, CLS and BRLi), reprocessed and distributed by Legos/CTOH, 2017).

2.3.4 Jason-1

Launched in December 2001 and ending in July 2013, the Jason 1 mission was a joint CNES and NASA initiative with the objective of monitoring ocean circulation, linking atmospheric and ocean activity, improving climate predictions, and monitoring global ocean cycles such as El Nino. During much of its mission Jason-1 followed the orbit of the TOPEX/POSEIDON mission, which would then be followed by Jason-2. This was a non-sun-synchronous ~10 day repeat orbit with an inclination of 66°. Jason-1 was equipped with the Poseidon-2 radar altimeter (5.3GHz (C-band) and 13.575 (Ku-band)). The onboard radiometer was the Jason Microwave Radiometer (JMR), a passive radiometer measuring at 18.7, 23.8, and 34 GHz. It used Doppler Orthography and Radio positioning and Integrated by Satellite (DORIS), LRA, and a GPS receiver know as Blackjack for positioning. Available data products include Sea Surface topography, Significant wave height, Surface wind speed, Ocean Tides, vertically integrated atmospheric water vapor and vertically integrated ionosphere electron content. They are available at 1 and 20hz(Cole et al., 2013; "Jason-1 Products Handbook," 2016; Kramer, 2002; Lafon & Parisot, 1998).

2.3.5 Jason-3

Launched in January of 2016 as a joint initiative between the National Oceanic and Atmospheric Administration (NOAA), Eumetsat, NASA, and CNES, the Jason-3 mission is the contemporary follow on to the TOPEX/Jason series of satellites following the same orbit. The mission hardware is quite similar to that of Jason-2. This includes a Poseidon (3b) altimeter, Advanced Microwave Radiometer (AMR), DORIS, LRA, and a GPS receiver (GPSP), as well as an experimental radiation detector (CARMEN-3) and light particle telescope (LPT)(Dumont et al., 2016).

Table 4 GDR Editing Criteria for Altimetry data

TOPEX, JASON-1, JASON-2, Jason-3 GDR-D data	ERS-1, ERS-2, ENVISAT GDR, Version 2.1 or later
Conditions of 1-Hz data parameters	
<ul style="list-style-type: none"> • Flag orbit_state_flag_rest = 3 (adjusted preliminary/precise orbit) • Range correction model_dry_tropo_corr is available • Range correction model_wet_tropo_corr is available • Range correction iono_corr_gim_ku is available • Range correction solid_earth_tide is available • Range correction pole_tide is available 	<ul style="list-style-type: none"> • Bits 30-31 of meas_conf_data_flags = 00 (adjusted preliminary/precise DORIS orbit) • Range correction mod_dry_tropo_corr is available • Range correction mod_wet_tropo_corr is available • Ku band range correction of GIM ionosphere (not in Offline GDR) is available • Range correction solid_earth_tide_ht is available • Range correction geocen_pole_tide_ht is available
Conditions of 20-Hz data parameters	Conditions of 18-Hz data parameters
<ul style="list-style-type: none"> • Altitude alt_20hz is available • Ku band altimeter range (ice retracking) ice_range_20hz_ku is available • Flag ice_qual_flag_20hz_ku = 0 (good, Ku band ice retracking quality) • Ku band backscatter coefficient (ice retracking) ice_sig0_20hz_ku is available and not negative (in dB) • The number of 20-Hz data points to be compressed to 1-Hz in a selected pass segment has to be 2 or more. • The time span of each pass segment should not be longer than 1.5 seconds (this limit was removed in processing some segments exceptionally large virtual Stations). 	<ul style="list-style-type: none"> • Altitude hz18_diff_1Hz_alt is available • Flag bit of map_18hz_ku_trk_flags = 0 (valid, Ku band tracker range validity) • Ku band altimeter range hz18_ku_ice1 is available (ice1 retracking) • Flag bit of ku_ice1_retrk_qua_flags = 0 (good, Ku band ice1 retracking quality) • Ku band backscatter coefficient hz18_ku_ice1_bscat is available and not negative (in dB, ice1 retracking) • The number of 18-Hz data points to be compressed to 1-Hz in a selected pass segment has to be 2 or more. • The time span of each pass segment should not be longer than 1.5 seconds (this limit was removed in processing some segments exceptionally large virtual Stations).

2.4 Global River Widths from Landsat and RivWidthCloud

The Global River Widths from Landsat (GRWL) database is a 30-m resolution collection of river center lines and their associated widths at mean annual discharge, along with a braiding index. The dataset consists of >58 million river measurements with global coverage (Allen & Pavelsky, 2015). Mean discharge dates were determined for all observed rivers by analyzing mean monthly discharge values from the Global Runoff Data Center (GRDC). Widths and braiding index values were generated by creating binary water masks of ice- and cloud-free Landsat images and processing with RivWidth software (Allen & Pavelsky, 2015; Pavelsky & Smith, 2008).

The most contemporary version of the river width data extraction from Pavelsky et al., is RivWidthCloud (X. Yang et al., 2019). This software allows users to run a newer version of the original RivWidth that was used in the construction of GRAWL on Google Earth Engine (Pavelsky & Smith, 2008). We used RivWidthCloud to extract additional width data needed for generating Channel Water Storage data that is a part of GRRATS1kd derived products (section 3.5).

2.5 River Ice Data

Ice elevations can represent the water surface elevation when a river is partially frozen, but they decouple from water elevations when the river completely freezes. GRRATS includes an ice cover flag rather than height data when rivers are partially frozen. Rivers that are frequently ice covered include the Indigirka, Kolyma, Lena, Mackenzie, Mezen, Ob', Olenyok, Pechora, Yenisei, and Yukon rivers. Wherever possible we used MODIS ice cover detection to determine if rivers were frozen. Based on our evaluation of Landsat imagery, the St Lawrence River has not completely frozen over during the period covered by GRRATS. Public information is available for North American river freeze-up dates from the United States Geological Survey (USGS) and Environment and Climate Change Canada (ECCC). For the Yukon River, ice dates were gathered using USGS stream gage 15565447; USGS includes in their stream gage data a seasonal operation flag that indicates if the stages are affected by or missing due to freeze-up (USGS, 2016). Ice dates for the Mackenzie River were based on a ECCC gage (10LD001) with similar seasonal operation (Environment Canada, 2016). As there is no modern information publicly available about ice conditions of the Asian Arctic rivers in the dataset, we obtained climatological values from *Pavelsky and Smith* (2004), establishing general temporal limits for ice cover in the region. Remote sensing of river ice cover typically requires a combination of sensors and ultimately some kind of visual verification. It would be ideal to use radar backscatter from the radar altimeters used in this study, but this metric can vary greatly with air inclusions in the ice and surficial meltwater (Jeffries et al., 2013). Our efforts to detect ice cover were inconsistent.

2.6 The Shuttle Radar Topography Mission (SRTM)

The SRTM was a joint effort between NASA Jet Propulsion Laboratory (JPL) and the National Geospatial-Intelligence Agency (NGA) which used interferometric Synthetic Aperture Radar (InSAR) to generate a globally-consistent digital elevation model (DEM) for latitudes less than 60°. With 30-m resolution, SRTM data are an order of magnitude higher resolution than predecessors such as GLOBE and MOMS-2p (Hastings & Dunbar, 1998; Schroeder et al., 2000). The 10-day shuttle-borne mission in February 2000 succeeded in creating a three dimensional image of 80% of the earth's land surface (Van Zyl, 2001). NASA and the NGA conducted a considerable ground campaign to validate SRTM, allowing the estimation of height uncertainties, throughout the globe, that accompany the data product. The

absolute height and geolocation error are <10 m for the majority of the earth and range from 2.6 m to 12.6 m (Rodriguez et al., 2006).

2.7 GMTED2010

The Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) is the result of collaborative efforts between the USGS, the Earth Resources Observation and Science Data Center (EROS) and the National Geospatial-Intelligence Agency (NGA) to improve on the existing Global 30 Arc-Second Elevation (GTOPO30) DEM. (Danielson & Gesch, 2008, 2011). GTOPO 30 was compiled using data primarily from Digital Terrain Elevation Data (DTED) and Digital Chart of the World (DCW) combined with several smaller datasets with spot height data (USGS-Eros data center, 1997).

GMTED2010 was generated in three different resolutions (1k m, 500 m, 250 m) based on input data resolution and comprised a collection of raster products for each resolution including minimum elevation, maximum elevation, mean elevation, median elevation, and standard deviation of elevation. Approximately 70% of the data used for height corrections and improved spatial resolution in GTOPO30 were derived from SRTM data, allowing GMTED2010 to exhibit better topographic detail than its predecessor (Carabajal et al., 2011).

2.8 ASTER

The ASTER sensor, onboard NASA's Terra satellite produces imagery in both visible and near infrared (VNIR) bands at a 15 m spatial resolution, as well as short wave and thermal infrared bands at 30 m and 90 m resolutions, respectively. Computing the ASTER DEM required VNIR imagery via photogrammetry, enabled by a 15 m along-track stereo-band looking 27.6° backwards from nadir. The ASTER DEM thus has global coverage (Abrams, 2000), but also has well-known height errors and has frequently shown RMS errors far higher than SRTM (Huggel et al., 2008).

2.9 Validation gages

We used publicly accessible river stage gage data for the validation of GRRATS whenever possible. In the United States, we compiled data from the United States Geological Survey (USGS) records (USGS, 2016). We sourced Canadian data from the ECCC Water Survey of Canada (WSC). We acquired gage data for South America through SO HYBAM. Data for the Mekong River is available through the Mekong River Commission (MRC), where records from 2008 and onward are available for the wet season. We acquired data for Africa, India and Bangladesh (data from the Bangladesh Water Development Board BWDB), through information sharing with collaborators, meaning it is not publicly available. We listed the source of validation data for each river in Table 5: Validation gage information.

GRRATS is a global river elevation dataset with time series from 902 locations on the world's largest rivers, complete with a full set of production metadata (including the unprocessed level -2 data) and QA/QC for each VS (Table 9-Table 13). It resulted from the assessment data extracted at all potential VS locations, located on ocean-draining rivers >900 m wide (sections 3.1 and 3.2). We utilized a unique quality filtering process based on a priori data to remove erroneous measurements and historical data to flag times of seasonal ice cover (section 3.3).

Table 5: Validation gage information

River	Gage ID	agency/source
Amazon	10099800, 13155000, 14050000, 14100000, 15040000, 16030000,17050002	SO HYBAM
Brahmaputra	Aricha, Bahadurabad_Transit, Chilmari, Fulchari, Jaganathganj, Kamarjani, Kholabarichar, Mathura, Mathurpara-Milanpur, Noonkhawa, Porabari, Sirajganj	BWDB
Columbia	12472800,14128600, 14128870, 14144700, 14246900	USGS
Congo	Brazaville, Kindu, Kinshasa, Kisangani, Lisala	S.Calmant
Ganges	Hardinge Bridge, Mohendrapur, Panka, Rajsjahi, Sardah, Sengram, Talbaria	BWDB
Mackenzie	10FB001, 10KA001, 10LD001	WSC
Mekong	ChiangKhan, Kratie, Mukdahan, NakhonPhanom, NongKhai, Paksane, Pakse, PhnomPenh, Savannakhet, StrungTreng TanChau, Thakhek, Vientiane,	MRC
Mississippi	5378500, 7010000, 7374000	USGS
Orinoco	40800000	SO HYBAM
St Lawrence	02MC005, 02OA016, 02OB011	WSC
Susquehanna	1554000, 1570500, 1576000	USGS
Yukon	15356000, 15453500, 15565447	USGS

3. Methods

GRRATS is a global river elevation dataset with time series from 902 locations on the world’s largest rivers, complete with a full set of production metadata (including the unprocessed level -2 data) and QA/QC for each VS (Table 8-Table 12). It resulted from the assessment data extracted at all potential VS locations, located on ocean-draining rivers >900 m wide (sections 3.1 and 3.2). We utilized a unique quality filtering process based on a priori data to remove erroneous measurements and historical data to flag times of seasonal ice cover (section 3.3).

Wherever possible we compared VS with in-situ gages. In lieu of gage data, we visually inspected each time series and gave a qualitative assessment (section 3.4). Figure 1 shows the entire data extraction process as a flow chart.

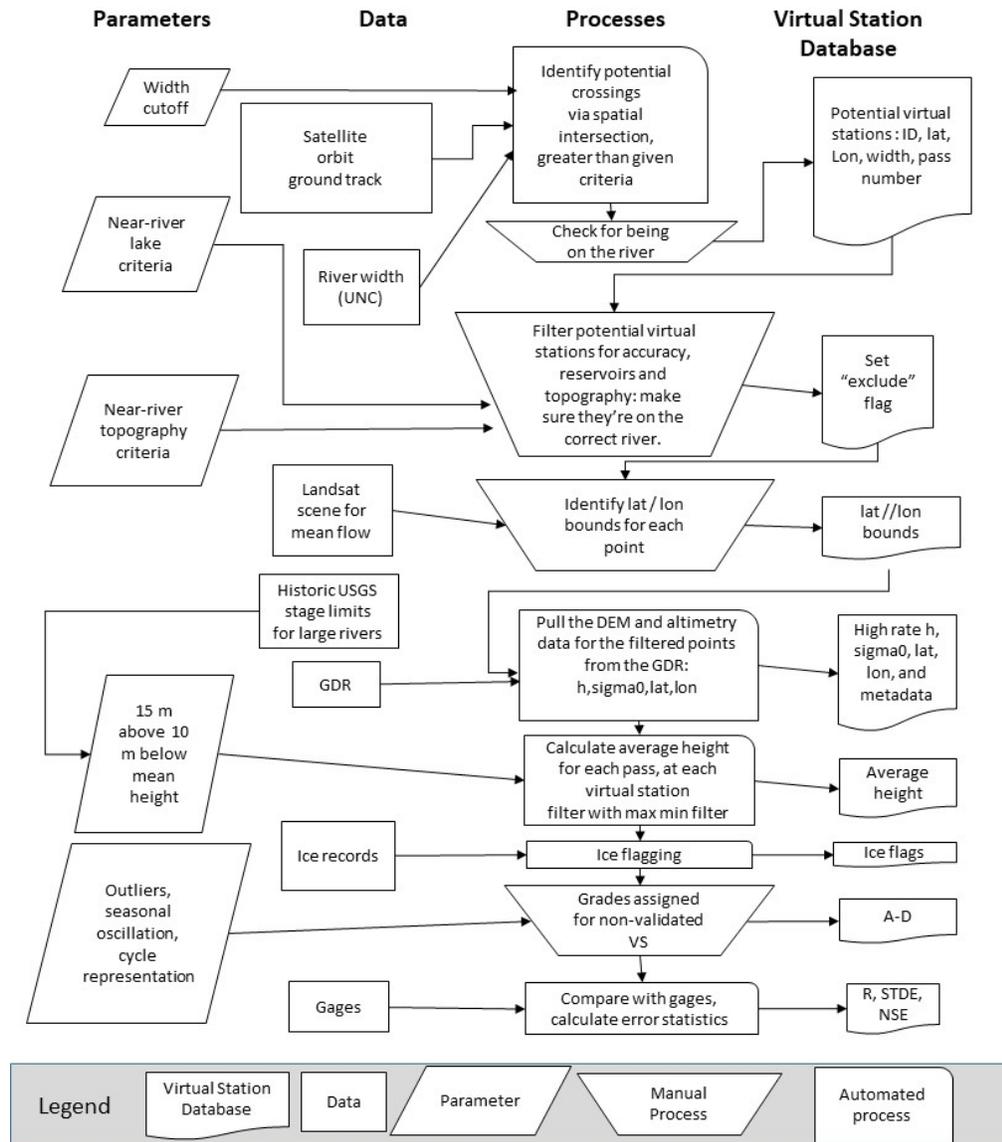


Figure 1 Data Processing flowchart for GRRATS.

3.1 Potential VS Selection

We identify potential VS by overlaying all ocean-draining river centerlines from GRWL with width greater than or equal to 900m onto Landsat imagery. The 900 m width cutoff is in part a pragmatic decision; in general, we find that returns from rivers smaller than 900 m were less likely to produce usable results. We then find intersections between filtered centerlines and ENIVSAT and the OSTM/ Jason-2 ground

tracks, which included a 2 km buffer zone measured perpendicularly to the ground tracks to accommodate deviations from the nominal orbit (Beckley et al., 2013). While it is true that rivers shift over time, planform migration is inversely correlated with width (Constantine et al., 2014; Hickin & Nanson, 1984). We can expect minimal impact from migration on rivers 900m or wider.

3.2 Extraction of altimetry data

We overlay mean discharge Landsat imagery compiled for the GRWL river centerlines dataset (Allen & Pavelsky, 2015) with ground tracks from the included radar altimeter satellites to define the edges of the mask used for the extraction of water elevations. We selected these images because they are temporally representative of median discharge and had very little in the way of cloud cover (Allen & Pavelsky, 2014). We manually draw polygons at each potential VS, and trace the river boundary on the GRWL mean discharge Landsat image; the ± 2 km buffer zone perpendicular to altimeter ground track is taken into account in each the polygon shape, in order to allow for orbital drift (Figure 2). On average, there are only a handful of individual satellite returns over the river, for each pass. For example, Jason-2 ground speed is 5.8 km/s, and the high-rate channel measures at 20 Hz; thus, there will be on average 2.5 samples located within the polygon for a 900 m wide river. We extract all altimeter returns falling within each polygon for each pass from Jason-2 Geophysical Data Record (GDR) version D (Dumont et al., 2009), and the Envisat GDR, Version 2.1 or later (Soussi & Féménias, 2009), using corrections outlined in Table 4 GDR Editing Criteria for Altimetry data. We extract Ice-1 retracked heights from the GDR. We then convert these ellipsoidal heights to a height above the geoid, using the EGM08 model

(Pavlis et al., 2012)



Figure 2 A sample Landsat Scene from the Amazon River with to VS polygons and their overlying Envisat ground tracks.

Waveform retracking in the literature, when described, is often applied in customized ways for each river as in Schwatke et al. (2015) and Berry et al. (2005). Work such as Boergens et al. (2016) has demonstrated new retracking methods that are not customized to each river, but none have shown that their new approach has been tested outside of a case study. Ice-1 performance is robust and often the best choice for consistency across a basin with varying widths (Biancamaria et al., 2017b). We have elected to use Ice-1 only, as opposed to custom retracking for each VS because custom retracking would require a departure from using uniform methods across the entire dataset. The Ice-1 retracker is based on the Wingam et al. (1986) Offset Center of Gravity Retracker and is described in further detail in Gommenginger et al. (2011). Ice-1 has consistently been shown to be a robust retracking algorithm for rivers (Biancamaria et al., 2017b; Frappart et al., 2006; Santos da Silva et al., 2010).

3.3 Filtering of altimetry data

We identified four primary issues in our initial data exploration, which our algorithm addresses: 1. Near river topography frequently causes on board tracking errors resulting in height distributions where statistically distinguishing the correct measurements is difficult without other information. 2. Height decay as heights approach base flow are often erroneously low. 3. Ice cover can result in measurements

that are not representative of the river surface. 4. Data sparsity limits the utility of a time series after removal of bad data points. Therefore, we base inclusion in the final product on four filter criteria: 1. Individual data point removal (one altimeter cycle) based on expected height range from a priori data, 2. Individual data point removal of erroneously low values and 3. Individual data point removal during times of ice cover, 4. time series removal due to limited data volume (the entire VS is omitted from the final product).

Rather than using a purely statistical method i.e., the along track outlier test used for Schwartke et al. (2015) to filter altimetry returns, we filter GRRATS river surface elevations based on a priori DEM data from inside the extraction polygon (criteria 1). Readers can find more information on what methods each dataset uses, in its reference in section 2. We filter using a predefined baseline elevation at each VS. We extract an initial estimate of the baseline from the best available DEM within each VS polygon: SRTM, GMTED2010 or ASTER, in order of priority. This priority is determined by assessing the consistency of each DEM over long-profiles of several of the largest rivers. We compute the initial baseline as the median of all best available DEM values falling within the polygon for each VS. Each VS location was associated with a flow distance based on the river center lines from the Hydro1k dataset (EROS, 1996). We then use a simple linear optimization algorithm to force the initial baseline heights to monotonically decrease downstream, a step that is necessary due to noise in DEM data. The algorithm minimizes the absolute value of the difference between the constrained baseline and initial baseline elevations at each VS, subject to the constraint that the elevations monotonically decrease downstream. We construct the final height filter from these constrained baseline heights, referred to as “baseline” hereafter.

In order to establish the expected range of river surface elevations with respect to the baseline, we compute river elevation variations from in-situ gages. We examined over 150 USGS gages for watersheds larger than 20,000 km² in size, and found that height filter limits of 15 m above or 10 m below the constrained baseline height should reasonably encompass any measurements of the river surface. We then apply a second filter to altimeter returns that remained to remove potentially erroneous returns on the low end of that data (criteria 2). Any heights that fall under 2 m below the 5th percentile of height are removed (0.03% of total returns). We obtained these filter criteria by trial-and-error, and found it to be necessary to remove elevation observations excessively impacted by near-river topography at low flow.

Many rivers in the database experience seasonal ice cover (criteria 3). Thus, we include the observations of the unfiltered river heights in the GDR, but not in the filtered time series. River ice cover dates were determined from USGS and WSC data when available. A USGS station near the mouth of the river was selected, as breakup typically occurs last at the river mouth (Cooley & Pavelsky, 2016). When freeze and thaw data are not directly available (e.g. for Siberia), we applied broad date limits regionally, using observations from the *Pavelsky and Smith (2004)* study of Arctic river ice breakup.

Both the unfiltered and filtered elevations are included in the GRRATS data files. After height filtering, we remove any potential VS that was missing 50% or more altimeter cycles from the final VS data product (criteria 4). VS are required to contain 25% of the original cycle data after ice filtering in order to remain in the final dataset. The number of virtual stations in each basin are listed in Table 3 and shown in map view in Figure 3

In GRRATS V2 a new set of statistically filtered heights is introduced. Following a modified version of the methods outlined Tourian et al.,(2016), we took all available height data from all altimeters on the river, rescaled between 0 and 1, and then used a moving window (scaled to average repeat interval) to flag and remove values outside of 95% confidence interval (± 2 m). These measurements were removed before pass averaging in the SF height data.



Figure 3 VS and validation sitemap. Red circles show locations of 9020 virtual stations. Black triangles show the locations of 50 validation gauge locations.

3.4 Validation

We acquired validation stage data from 65 globally distributed in-situ stream gages detailed in Table 6 for 12 of the 39 rivers in the data set. These 12 rivers with at least one gage include 380 of the 920 virtual stations, or approximately 42% of the VS in GRRATS. Note that VS rarely fall in the same location as a stream gage; thus, most studies recommend some VS-in-situ stream gage distance (e.g. 200 km) beyond which comparisons are not performed (Michailovsky et al., 2012). However, our preliminary analyses showed that VS-stream gage distance was often not an accurate predictor of height anomaly differences. For example, Figure 5 shows all VS on the Orinoco and an in-situ stream gage prior to mean difference removal. The lack of correlation between VS-stream gage distance and height anomalies certainly warrants further investigation, but it is outside the scope of this study. There is great similarity in the time series from VS that are often great distances apart; Figure 6 shows the spatial distribution of altimeter crossings for the Orinoco. Thus, in this study, we compare each virtual station with each in-situ gage available for that river. At each VS, we report error metrics for both the best, median, and the closest comparison. For completeness, we include VS with poor error metrics; users can then select which of the VS to use, based on their reported error statistics.

Following Schwatke et al. (2015), we compare relative heights between gages, as opposed to absolute heights, in order to avoid the influence of datum errors and the lack of correspondence between satellite ground tracks and gage locations. We calculate relative heights by removing the long-term

mean of difference between the sample pairs of VS heights and the stage measured by the stream gages. Principles of hydraulic geometry dictate that the most precise comparisons would account for change in cross sectional geometry (affecting STDE) and time lag (affecting which gage data points were compared to the VS) between points of comparison (C. J. Gleason & Smith, 2014; Leopold & Maddock, 1953). However, the primary purpose of our validation was to provide mass evaluation of the dataset to aid end users in selection of VS for their work; therefore, a more simplistic validation approach was

Table 6 Identification and source of validation gages in present study

judged adequate. Error metrics included in GRRATS include the correlation coefficient (R), Nash-Sutcliffe

River	Gage ID	agency/source
Amazon	10099800, 13155000, 14050000, 14100000, 15040000, 16030000,17050002	SO HYBAM
Brahmaputra	Aricha, Bahadurabad_Transit, Chilhari, Fulchari, Jaganathganj, Kamarjani, Kholabarichar, Mathura, Mathurpara-Milanpur, Noonkhawa, Porabari, Sirajganj	BWDB
Columbia	12472800,14128600, 14128870, 14144700, 14246900	USGS
Congo	Brazaville, Kindu, Kinshasa, Kisangani, Lisala	S.Calmant
Ganges	Hardinge Bridge, Mohendrapur, Panka, Rajsjahi, Sardah, Sengram, Talbaria	BWDB
Mackenzie	10FB001, 10KA001, 10LD001	WSC
Mekong	ChiangKhan, Kratie, Mukdahan, NakhonPhanom, NongKhai, Paksane, Pakse, PhnomPenh, Savannakhet, StrungTreng TanChau, Thakhek, Vientiane,	MRC
Mississippi	5378500, 7010000, 7374000	USGS
Orinoco	40800000	SO HYBAM
St Lawrence	02MC005, 02OA016, 02OB011	WSC
Susquehanna	1554000, 1570500, 1576000	USGS
Yukon	15356000, 15453500, 15565447	USGS

Efficiency (NSE), and Standard deviation of the errors (STDE). NSE is a fit statistic that is typically employed to describe the goodness of fit for a modeled result with measured values, so our use here is non-traditional. We have chosen to use NSE, because it has value as a fit statistic, which goes beyond R and STDE alone. NSE normalizes error with variation from the mean in the observed, or in our case, in-situ data. It is an established metric for goodness of fit within the altimeter literature (Biancamaria et al., 2018; Tourian et al., 2016). For 27 of the 39 rivers in the GRRATS dataset, no in-situ data is available for evaluation. For these stations, we perform a qualitative evaluation of the station represented by a letter grade ranging from A (highest level of confidence on the data quality) to D (lowest level of confidence).

The criteria used in the assignment of letter grades was based on the presence of obvious outliers, number of data points in the time series, and time series continuity with nearby VS. We determined outliers by visual inspection. While qualitative grades are not as reproducible as fit statistics, they have

been used in the past to guide users to preferable time series when no other error metrics are available (Birkett et al., 2002).

3.5 GRRATS1kd

GRRATS1kd is a derived 1km daily resolution interpolated product with surface heights and channel water storage (CWS) measurements. Our interpolation algorithm is based on bi-linear solution Delaunay triangulation over a space (1 dimensional) time grid of height values. Each VS has its time series mean removed, prior to interpolation. We separate height data into different groups based on concurrent mission timing before interpolation. We interpolate then combine the resulting interpolation blocks to create a matrix of the entire sampling area over the entire sampling period. We then smooth these anomalies using a penalized least squares approach, with iterative outlier rejection (Garcia, 2010). We then add back a mean river surface height for each location based on SRTM, GMTED2010, or ASTER data in that preferential order. Finally, we force data from each time to go downhill in the direction of the mouth of the river using the same linear optimization approach used to create the base line heights I GRRATS.

For the width component of CWS, we collected width measurements using RivWidthCloud (X. Yang et al., 2019), a python tool that uses the google earth engine to classify Landsat scenes, assign channel centerlines and measure width orthogonally to those center lines at 30m resolution. We assigned points to their nearest 1km section and took a mean value for each section/time combination. We generated missing widths by fitting a piecewise linear regression to observed data where there were at least 10 observation pairs, and correlation of at least 0.5. In cases where these criteria were not met, we used a static mean width value for CWS calculation.

Finally, we calculated CWS by using width and height observations to estimate cross-sectional area for each daily height/width value using the method described in Durand et al., (2014). This cross-sectional area multiplied by the length of river it represented (always 1km for our product), to generate CWS in km^3 . Further process details and the code used to calculate Cross-sectional area can be found at (<https://github.com/mikedurand/SWOTAprimeCalcs>).

4. Results and discussion

4.1 Filtering and VS selection

We extracted total of 2,844,704 individual radar returns at 1478 potential VS from the GDR. Of these, we found 802,438 (56.6 %) to be viable after height and ice filtering. Averaging all filtered returns within the river polygons for each pass at each VS led to a total of 181,723 pass-averaged measurements. VS were required to retain 50% (without ice) or 25% (with ice) of their passes post-filtering to be included in the final data product. This integrity requirement resulted in the removal of 524 potential VS locations. VS were also removed by visual inspection if they were tidal, resulting in the removal of an additional 46 stations. For the VS included in GRRATS, 68.5% of the total returns passed the filters. The median percentage of lost returns among the VS included in GRRATS was 2.9%. There were 224 VS that lost no returns during the filtering process. Almost 60% of the VS (524 of 917) lost fewer than 10 returns

during filtering.

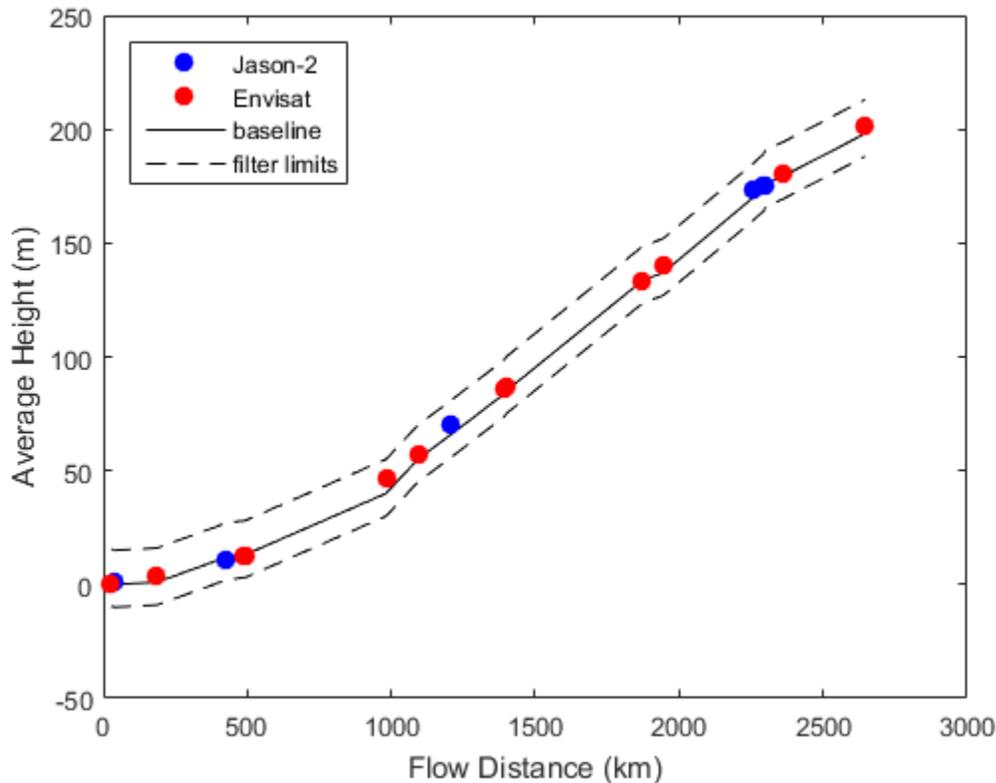


Figure 4 Example filter limits and average VS time series heights for the Mississippi River. The SRTM-derived baseline height represents the adjusted median of the returned heights inside the domain of each virtual station along the river. The filter thresholds are assigned to be 15 m above and 10 m below the baseline height. Circles represent the locations of Envisat and Jason-2 virtual stations along the Mississippi river.

Figure 4 shows good agreement between the DEM baseline height filter limits, and VS time series height averages as a function of flow distance along the Mississippi River. There are a total of 18 VS (6 from Jason-2 and 12 from Envisat) along the Mississippi. The mean difference between baseline and time series average altimeter height is 1.66 m, with a standard deviation of 2.01 m. A high bias is consistent with the fact that most erroneous altimeter returns result when height measurements are dominated by nearby topography rather than the river surface (Frappart et al., 2006; Maillard et al., 2015; Santos da Silva et al., 2010). On average, Jason-2 returns are 1.37 m higher than the baseline and Envisat returns are 1.8 m higher than the baseline; this result reveals a mean bias between the two altimeters of ~ 43 cm. This bias is likely due to issues with modeling of ionic and troposphere delay and can occur on a decimeter level (Fernandes et al., 2014), as well as differences in footprint size, altimeter power, and other characteristics of the sensors and algorithms.

Space-time coverage varies significantly among rivers, depending upon the relative locations of altimeter ground tracks and river orientations. For example, the gap in the time series on the Orinoco River after the transition from Envisat to Jason-2 time series that occurred in ~ 2008 is apparent in Figure 5.

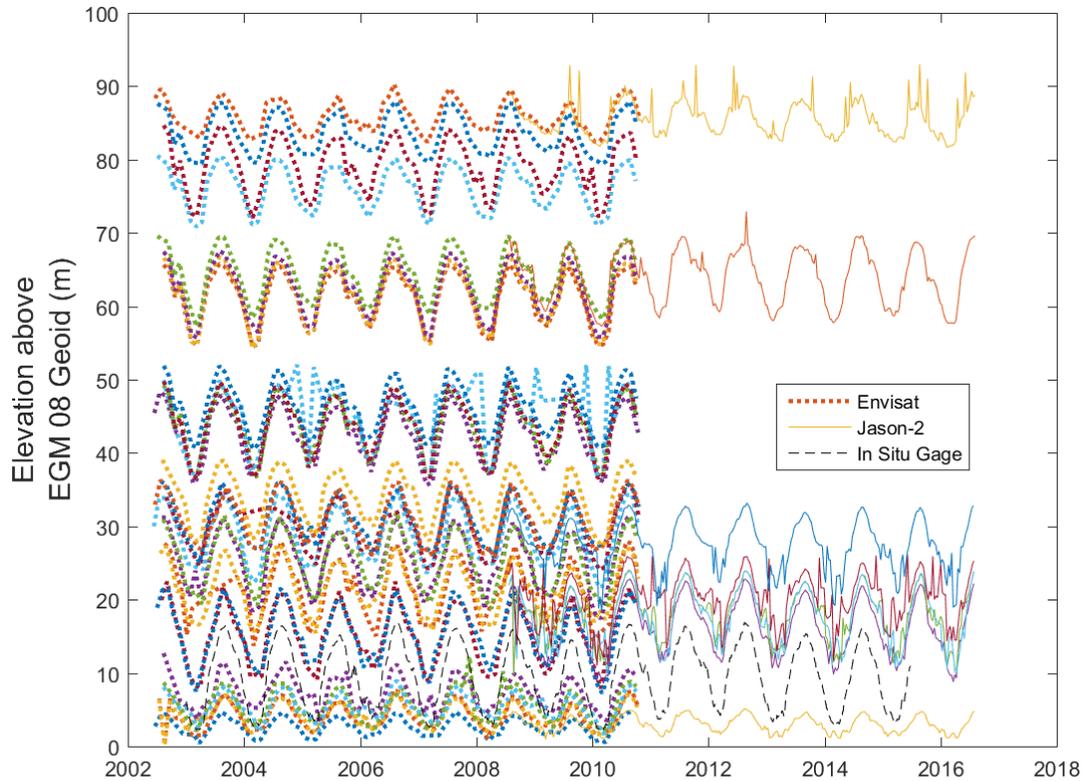


Figure 5 All VS time series on the Orinoco River and validation gage (dashed line) prior to mean difference removal. Envisat and Jason-2 data are plotted as dotted, and solid lines, respectively. The sparseness of data in later dates of the composite are due to the intersection of the river with the different mission ground tracks and can be seen in map view in Figure 6.

A map view of the region (Figure 6) shows that the Jason-2 ground track (red) is much less dense than that of Envisat (yellow). The river flows northward for a long stretch nearly parallel to the Jason-2 ground track, and is thus not well-sampled, but it is well sampled by Envisat. Similarly, the cluster of VS below the gap in Figure 5 is due to the river course being essentially overlapped by the Jason-2 ground track in this location.

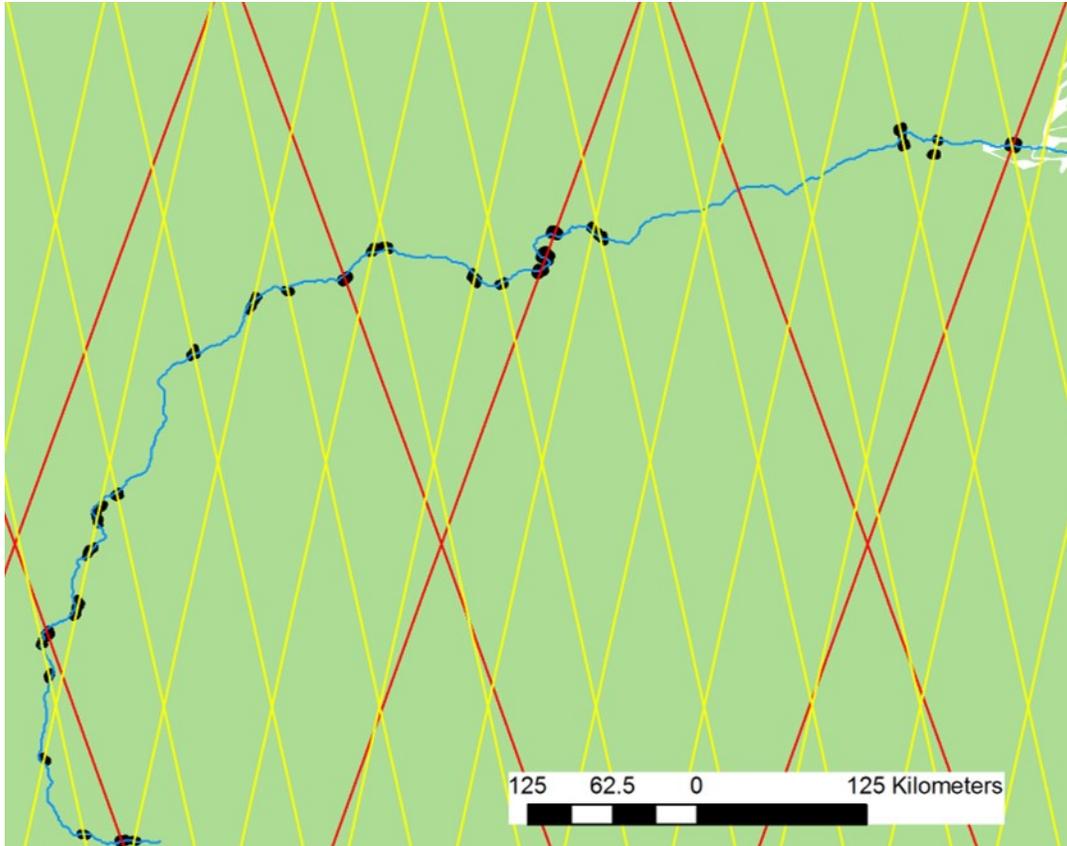


Figure 6 Blue line shows the centerline of the Orinoco River. The black circles indicate the location of the virtual stations. Red lines represent the Jason-2 ground track, whereas the yellow lines show the Envisat ground track. There is a high number of Envisat ground tracks that intersect the river near its upper reaches (southwest) while the river course runs nearly parallel to the Jason-2 ground track. This causes a gap in the composite of all the time series in these reaches after the end of the Envisat mission.

4.2 Example time series evaluation

Figure 7 is an example GRRATS time series for the Mackenzie and Amazon Rivers and corresponding in-situ gages. Comparison between the Jason-2 time series and the gage on the Mackenzie River produced $STDE = 0.5$ m, $NSE = 0.41$, and an $R = 0.64$. In this case, the gage used for validation was located ~ 700 km upriver (Figure 7A). The STDE is approximately consistent with what is expected from the literature (Asadzadeh Jarihani et al., 2013; Frappart et al., 2006). However, the STDE is relatively large in comparison with the overall annual range in the time series (typically ~ 2 m) observed from the gage (see Figure 7A), leading to a relatively low NSE. Additionally, several passes have far larger errors, reaching up to two meters, in some cases. There were a total of 3 in-situ gages on the Mackenzie River. Across the 3 comparisons, this VS had median statistics of 0.58 m, 0.35 and 0.64 for STDE, NSE, and R, respectively.

Comparing the VS data to the gage on the Amazon River yields $STDE = 0.98$ m, $NSE = 0.94$ and $R = 0.97$, with the validation gage 263 km upriver from the VS (Figure 7B). Despite the STDE being nearly twice as large, the magnitude of change on the Amazon allowed for a much better fit due to the large interannual variability of the Amazon floodwave (>10 m). Most of the error is from times of low flow

near the ends of the calendar year in 2009, 2011 and 2012. There are 6 in-situ gages on the Amazon River. Across these comparisons, this VS had median statistics of 0.94 m, 0.95, and 0.98 for STDE, NSE, and R, respectively.

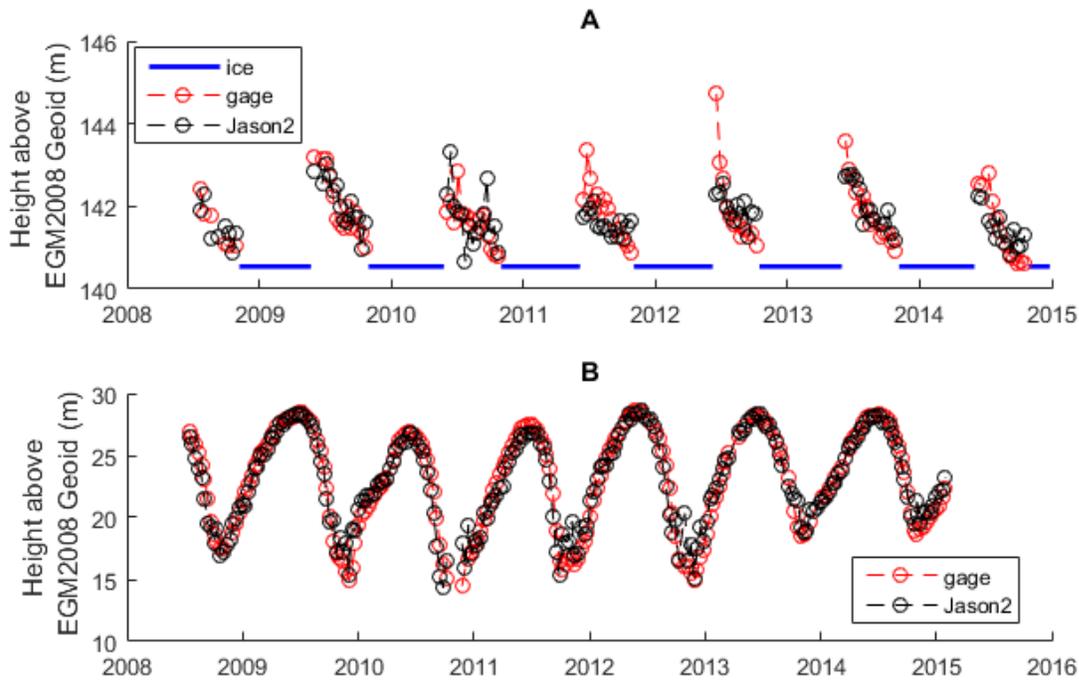


Figure 7 Example time series for the Mackenzie River. Panel A shows water surface heights with ice filtering compared to WSC gage (10KA001) located 684 km away from the virtual station. Panel B compares the time series derived from Jason-2 for one of the Amazon

As noted above, we used a novel strategy to compare VS and in-situ gages, comparing all VS to all gages on each river, regardless of spatial separation between the two. The relationship between the NSE and STDE error statistics for every VS-gage comparison on the Amazon river predominantly show decay in fit with distance (Figure 8A and B). There are 6 in-situ gages and 90 VS on the Amazon River; there are thus 630 comparisons. It is clear from Figure 8 that one would expect greater NSE and lower STDE at short VS-gage separation distances. On the other hand, performance at even the shortest gage-VS separations ranges from STDE of nearly zero to greater than 3.0 m, and NSE values from nearly 1.0 to less than 0.5. Similarly, stations nearly 1,000 km apart have less than 1.0 STDE, and greater than 0.75 NSE. From figure 8C and D, results from the Brahmaputra River, it is even more evident that this decay of fit over distances does not always hold and can change case by case. We found that in 41% of NSE, 41% of STDE, and 34% of R comparisons, the closest gage was not the one with the best-fit value. Thus, it seems somewhat limiting to impose a VS-gage separation distance beyond which we do not perform comparisons, as done in other studies. As noted above, we proceed by reporting for each VS the best and the closest comparison across all in-situ gages on a given river.

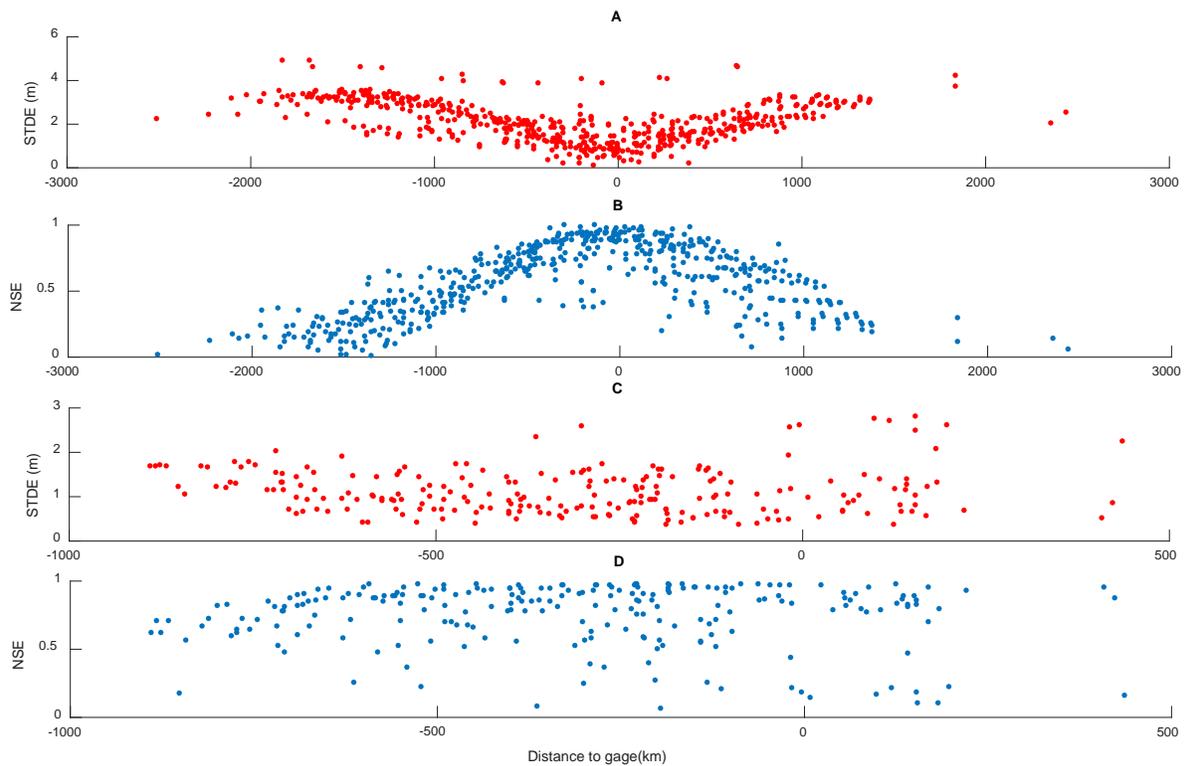


Figure 8 Error standard deviation (STDE) and Nash-Sutcliffe efficiency (NSE) of all the virtual stations on the Amazon River (A and B) and Brahmaputra River (C and D) plotted against gage proximity (gage location is at 0, negative distances are upriver). The Amazon has 90 VS that were compared with 7 gages, and the Brahmaputra has 19 VS that were compared with 12 gages.

4.3 GRRATS Evaluation across all rivers

There were in-situ validation data on a total of 12 rivers; on each river, the total number of time series evaluations is the product of the number of VS and the number of gages. Thus, the total number of time series evaluations (summed across all 12 rivers) was 1,894. For 32% of these time series evaluations, the VS was located within 200 km of in-situ gage. Figure 9 shows a map of the best NSE value at each VS with quantitative evaluation, and the letter grade for each VS with qualitative evaluation.

A total of 76.3% of the quantitatively validated virtual stations had an NSE greater than 0.4 when compared with at least one gage (Figure 10). The highest maximum NSE was 0.995, from an Envisat VS in the upper reaches of the Amazon. The median value for maximum NSEs for all VS was 0.8; note that the long tail in Figure 10A heavily influences the mean. A total of 332 of the 380 (87.37 %) virtual stations had a maximum NSE >0. The highest median NSE values were 0.96 at an Envisat VS in the mid Amazon, and two Envisat VS on the Orinoco river (lower and mid). A total of 296 of 380 (77.89%) had an median NSE >0. Positive values of NSE (Figure 10D) are approximately uniformly distributed between zero and one.

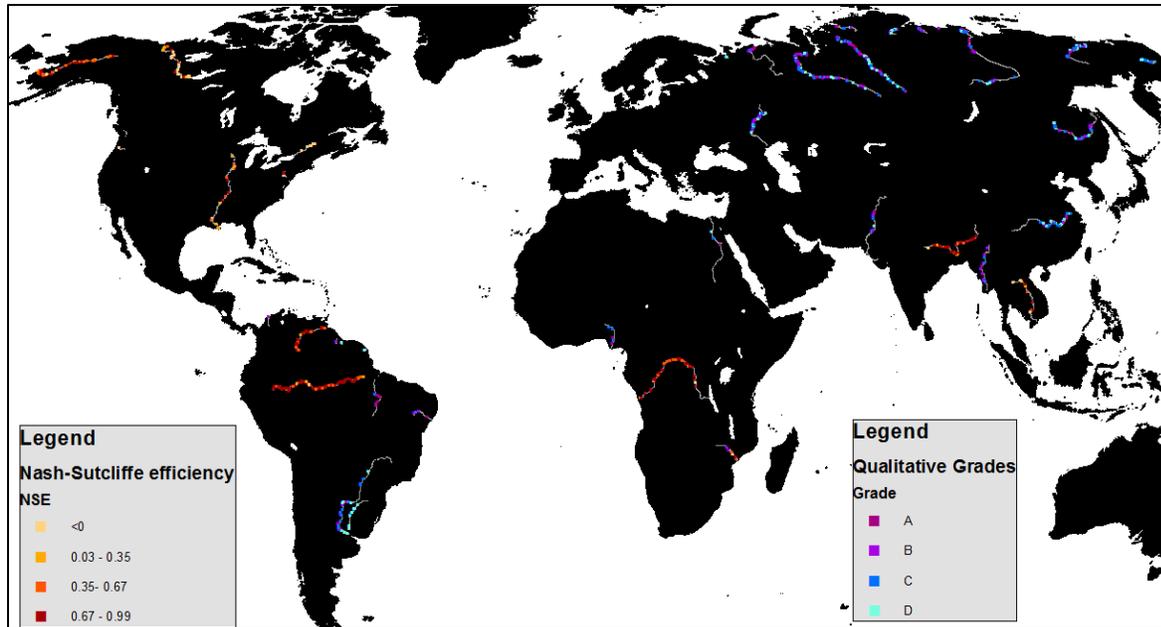


Figure 9 GRRATS validation results: Maximum NSE plotted in yellow to red and qualitative grades plotted in teal to dark purple. In both cases, darker colors indicate better validation results.

The smallest minimum STDE (to two significant digits) was 0.11 m and occurred at a Jason-2 VS on the lower Congo. The median value for minimum STDE for all VS was 0.84 m (Figure 10 A and C). The minimum and median value for median STDE were 0.31 m, and 1.28 m respectively. Our STDE error statistics are greater than previous work reporting accuracies ranging from 0.14 m to 0.43 m for Envisat returns and 0.19 m to 0.31 m for Jason-2 (Frappart et al., 2006; Kuo & Kao, 2011; Papa et al., 2012). This discrepancy is likely because GRRATS includes VS on rivers where validations have not previously been reported in the literature, and the fact that we do no fine-tune processing or filtering to each VS due to the global nature of the dataset.

Some locations with relatively low STDE values show poor performance in terms of NSE, particularly for rivers with relatively low water elevation variability. VS on the St Lawrence River had minimum STDE ranging from 0.58 - 3.27 m. The VS with 0.58 m STDE corresponded with a maximum NSE value of -0.27, indicating quite poor performance in resolving river variations (standard deviation of 0.35 m). The St Lawrence River is anomalous in other ways as well. For 2 potential VS (one each from Jason 2 and Envisat), the unprocessed data (Ice-1 retracked GDR data) showed a bias of several tens of meters above the baseline height, and thus no data for these VS are included in GRRATS. We double-checked the SRTM baseline elevations against other references for consistency. This case is particularly odd as such errors are not expected for wider rivers; the St Lawrence is between 2 and 7 km wide where we sampled it. Such errors are more commonly associated with altimeter returns from near-river topography on narrow rivers (Biancamaria et al., 2017a; Frappart et al., 2006; Maillard et al., 2015; Santos da Silva et al., 2010). There is also very little extreme topography or tall vegetation near the anomalous St Lawrence VS. It is unlikely that floating river ice contributed to these errors as the particularly problematic VS had poor performance throughout the year. Closer examination of these VS seems to indicate that the on-board tracking window was often 10s of meters outside of the river surface range, making retrievals from the surface impossible. Moderately poor performance from the remainder of VS

on the river is likely due to flow being under so many controls. The river lacks enough variation in height to allow for retrieval of a good signal outside the error range of radar altimeters.

The median of the maximum R values for each station is 0.92. The maximum R value plot (Figure 10B and D) shows left skewness, similar to the NSE results. The lowest maximum R value of -0.07 occurred at an Envisat station on the upper Mackenzie, which was the only virtual station to display a negative correlation. In addition to a poor correlation, the upper Mackenzie VS also showed poor performance in terms of the STDE (minimum STDE of 6.34 m) and NSE (max NSE of -20.84). Note that the Mackenzie is relatively deeply incised through much of its course, leading to a large potential for poor altimeter performance. The best maximum R value was 0.99 for an Envisat station on the upper Amazon that also displayed high NSE and low STDE. The median value of the average R is 0.66. The values range from -0.09 (an Envisat station on the lower St Lawrence) to 0.99 (an Envisat station on the lower Brahmaputra).

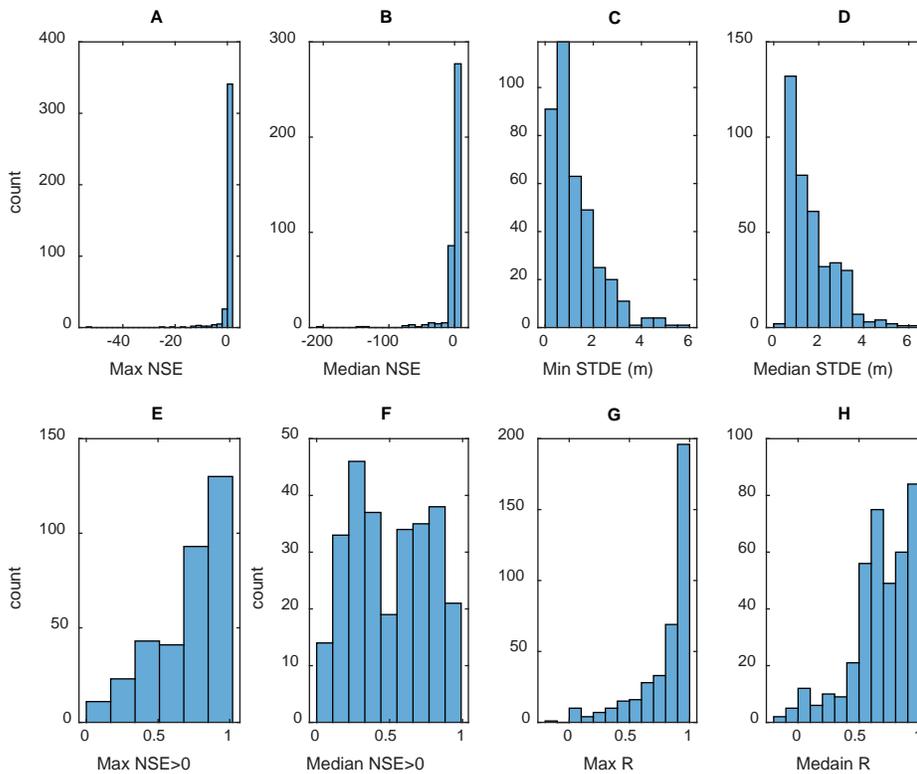


Figure 10 Virtual Stations Nash-Sutcliffe efficiencies computed with all available evaluation gages located in the same river. Panel A: histogram of the max NSE at each VS in the dataset, Panel B: histogram of the median NSE at each VS in the dataset, Panel C: histogram of the minimum STDE at each VS in the dataset, Panel D: A histogram of the median STDE in the dataset, Panel E: A histogram of the Max NSE at all the VS in the dataset with NSE>0, Panel F: histogram of the median NSE at all the VS in the dataset with NSE>0, Panel G: histogram of the max R at each VS in the dataset, H: histogram of the median R at each VS in the dataset.

We gave the remaining 27 rivers qualitative letter grades based on number of missing data points, obvious outliers, and agreement with nearby stations (Table 7). High to low grades were represented as A to D, respectively. To compute statistics across virtual stations with qualitative grades, A, B, C and D were represented as 4, 3, 2, and 1, respectively. Across all VS with qualitative grades, the median score

was 2 (C grade). There were 178 C grades representing ~33% of the sample. 85 (15.9%) VS received an A letter grade, while 114 (21.4%) received a D grade. Overall, the letter grades have a near normal distribution, though they do not pass a Chi-squared test at the 5% significance level.

Table 7 Grades from GRRATS VS that were evaluated qualitatively

Grade	A	B	C	B
Number of VS with grade	85	155	178	114

4.4 Towards quantitative performance prediction

As is evident above, radar altimeter performance varies dramatically across rivers and across VS. Generally, measurements from wide rivers without large topographic features in the altimeter footprints that have large seasonal water elevation variations tend to result in better altimeter performance. However, attempts to correlate GRRATS performance statistics to physiographic quantities (such as near-river topographic variability) were unsuccessful. The existing literature does not support a quantitative prediction of altimeter performance (Maillard et al., 2015). We computed average error statistics across all VS along each river (Table 8).

Table 8 River Average fit statistics

Best Average Statistics			Worst Average Statistics	
Fit Statistic	River	Value	River	value
Maximum NSE	Brahmaputra	0.82	St Lawrence Susquehanna Columbia Mackenzie	Max NSE<0
	Orinoco	0.78		
	Amazon	0.69		
	Ganges	0.65		
	Congo	0.6		
Maximum R	Orinoco	0.93	St Lawrence Mackenzie Columbia Susquehanna	0.3 0.46 0.49 0.68
	Brahmaputra	0.92		
	Ganges	0.87		
	Congo	0.85		
Minimum STDE	Congo	0.53 m	Mekong Orinoco Mackenzie St Lawrence	2.61 m 1.95 m 1.88 m 1.69 m
	Yukon	0.76 m		
	Brahmaputra	1.07 m		
	Mississippi	1.18 m		
Letter Grade		3.77	Courantyne Mezen Oiapoque	1.0
	Tocantins	3.27		
	Ayeyarwada	3.0		
	Sao Francisco	3.0		
	Olenyok	3.0		
Niger	3.0			
Essequibo	3.0			

Some rivers stood out as particularly good or poor performers, but no broad geographical patterns emerged. For this reason, we suggest that individual data point error be estimated as the STDE of its VS if it was validated. We recommend using the median value for validated STDE (0.84 m) as an error estimate for VS without validation data, as this is representative of 42% of all of the VS in the dataset. In order to identify conditions that may contribute to poor return quality, we compared both VS width and percentage of original returns post-filtering with all three fit statics. We found no statistically significant

relationships in this evaluation. Indeed, we are unable to explain the poor performance of many VS using width, near-river topography, or height variation. The St. Lawrence is an example of unexpectedly poor performance; typical predictors such as width (smallest VS ~1.5 km wide) and the lack of extreme proximal topography led to an expectation of accurate performance. All St Lawrence VS have demonstrated poor error statistics, as noted above. For one Jason-2 sampling location, a potential VS that was not included in the dataset near the mouth of the river, the GDR elevation was consistently 50 m average height, while true heights are close to sea level. The STDE for most VS is approximately 1 m (recall we compare height anomalies), but they are of limited use as the fluctuations in river height are also approximately 1 m.

Other rivers defy the normal pattern by showing good fit metrics while being far narrower. The Mississippi River was narrow enough that many radar altimeter crossings were excluded from the dataset. The VS widths range from 509.1 m to 2608 m, and had an average width of just 955.3 m. The average near-river relief ranged from 10-60 m. The Mississippi max NSE values ranged from -0.22 to 0.96 with an average of 0.43. Minimum STDE values ranged from 0.34 m-2.22 m, with an average of 1.18 m.

5.1 Data packaging and Variable identification GRRATS

Sample Altimetry Data (NetCDF format)

Format: netcdf4

Title = 'Altimetry Data for virtual station Yukon_Jason3_4'

Table 9 Global variables: The global variables are: longitude and latitude of the center of the virtual station, the virtual station ID, the satellite name, flow distance, sampling rate, the satellite pass number and a suite of fit statistics, or a qualitative letter grade. Qualitative letter grades were assigned based on amount of data points, seasonal pattern, and similarity to nearby VS. This was done, only when validation data was unavailable. When validation was possible, the VS was evaluated with all gauges

on the river through relative height comparison. Maximum Nash-Sutcliffe Efficiency (NSE), Average NSE, maximum R (correlation coefficient), minimum standard deviation of error (STDE), and average STDE are reported.

Variable	Dimension	Datatype	Units	Name
lon	X	double	degrees east	longitude
lat	Y	double	degrees north	latitude
ID	ID	char	-	Reference VS ID
sat	sat	char	-	satellite
Flow_Dist	distance	double	km	Distance from river mouth
rate	root	double	Hz	sampling rate
pass	root	int32	-	pass number
nse	grade	double	-	Max Nash Sutcliffe efficiency
nse median	grade	double	-	Median Nash Sutcliffe efficiency
R	grade	double	-	Correlation coefficient
std	grade	double	Meters	Minimum standard deviation of error
stdmedian	grade	double	Meters	Median standard deviation of error
prox	grade	double	Meters	The river centerline distance to the nearest gage.
proxSTD	grade	double	Meters	Nash Sutcliffe efficiency of most proximal gage
proxR	grade	double	-	correlation coefficient nearest
proxE	grade	double		Nash Sutcliffe efficiency of most proximal gage
nseSF	grade	double		MAX nash sutcliffe efficiency SF
nsemedianSF	grade	double		median nash sutcliffe efficiency SF
RSF	grade	double		correlation coefficient SF
stdSF	grade	double		MIN standard deviation of error SF
stdmedianSF	grade	double		median standard deviation of error SF
grade	grade	char	-	qualitative letter grade

Table 10 Groups: This includes the data from each return: lon and lat, the height of the water level in meters, the signal strength, sigma0, in decibels, a 'peakiness' value, the cycle number, the time of the return, and filter flags that signal 1 for data

that should be included and 0 for data that should be excluded. The flags are for a height filter, an ice filter, and the logical intersection of the two (allfilter), with 1 denoting returns that pass through the filter and 0 denoting returns that do not.

/Unprocessed GDR Data/				
Variable	Dimension	Datatype	Units	Name
lon	X	double	degrees east	Longitude
lat	Y	double	degrees north	latitude
h	Z	double	meters above EGM2008 geoid	Unprocessed heights
sig0	UGDR	double	dB	Sigma0
pk	UGDR	double	unknown	peakiness
cycle	UGDR	int32	unknown	Altimeter cycle
time	T	double	days since 1901-01-01T00:00:00	
heightfilter	UGDR	int32	-flag-	Good heights flag
icefilter	UGDR	int32	-flag-	No ice flag
allfilter	UGDR	int32	-flag-	Ice free heights that passed height filter

Figure 11 Table 11 Pass Averaged time series: These are pass-averaged values, having gone through the filter. There are two values that flag data: -9999 for data that is missing from the GDR, and -9998 for data that is missing because of height/ice filters. These flags are only present when none of the values to be averaged can be found. The other values give average height (hbar), in meters, and sigma-0 and peakiness weighted height.

/Timeseries/				
Variable	Dimension	Datatype	Units	Name
time	T	double	days since 1901-01-01T00:00:00	time
cycle	TS	int32	-	Altimeter cycle
hbar	Z	double	meters	average height
hwbar	Z	double	meters	weighted average height
sig0bar	time	double	dB	average sigma0
pkbar	time	double	-	Average peakiness
hbarST	Z	double	meters	Average_Height_above_EGM08_SF

Table 12 Sampling: This is data from the polygons, including the Landsat scene ID used to draw the polygons. The island flag is used when islands are visible inside the polygon in the imagery when drawing the mask.

/Sampling/				
Variable	Dimension	Datatype	Units	Name
scene	scene	char	-	Landsat Scene ID
lonbox	X	double	degrees east	Longitude box extent
latbox	Y	double	degrees north	Latitude box extent
island	scene	int32	-flag-	Island flag

Table 13 Filter data This is the filter data; nNODATA gives the number of cycles that have no data because of a lack of data in the GDR and/or data that is filtered out. riverh gives the river elevation extracted from a 30 arc-second DEM of the region. This is used for the height filter. maxh and minh are the upper and lower bounds of river heights included in the filtered data; we set a +15m, -10m from the DEM river elevation as a first pass, and then removed any data that was 5m below the 5th percentile of

river stage heights. icethaw and icefreeze are the thaw and freeze dates, respectively, for the years included in the altimetry dataset. DEM used refers to the DEM that the baseline height was taken from.

/Filter/				
Variable	Dimension	Datatype	Units	Name
nNODATA	-	int32	count	Number of cycles without data
riverh	Z	double	meters	River elevation from filter file
maxh	Z	double	meters	Max elevation allowed by filter
minh	Z	double	meters	Min elevation allowed by filter
icethaw	T	double	days since 1901-01-01T00:00:00	Thaw dates for river
icefreeze	T	double	days since 1901-01-01T00:00:00	Freeze dates for river
DEMused	DEM	Char	-	DEM used in height filter

5.2 Data packaging and Variable identification GRRATS1kd

Format: netcdf4

Title = 'Mississippi1kmdaily.nc

Table 14 Data in GRRATS1kd: Lat, and Lon describe the centroid of each 1km section of river. FDis the distance from the mouth of the river for each location. Time is the date of each event. Height is a time/flow distance grid of interpolated heights .Sat describes the active constellation of altimeters used for each times interpolation block. Storage is a time/ flow distance grid of CWS values.

Variable	Dimensions	Datatype	Units	Name
Lon	X	Double	degrees_east	longitude
lat	Y	Double	degrees_north	Latitude
Fd	Distance	Double	km	distance_from_river_mouth
Time	Time	Double	days since 1901-01-01T00:00:00	time
height	Time,distance	Double	m	interpolated_heights
Sat	Time,charlength	char	-	satellite
storage	Time,distance	Double	Km^3	river_channel_storage

6. References

- Abrams, M. (2000). The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER): data products for the high spatial resolution imager on NASA's Terra platform. *International Journal of Remote Sensing*, 21(5), 847–859.
- Allen, G. H., & Pavelsky, T. M. (2014). Patterns of river width and surface area newly revealed by the satellite-derived North American River Width (NARWidth) dataset (Vol. 1, p. 07). Presented at the AGU Fall Meeting Abstracts.
- Allen, G. H., & Pavelsky, T. M. (2015). Characterizing worldwide patterns of fluvial geomorphology and hydrology with the Global River Widths from Landsat (GRWL) database. Presented at the AGU Fall Meeting Abstracts.
- Asadzadeh Jarihani, A., Callow, J. N., Johansen, K., & Gouweleeuw, B. (2013). Evaluation of multiple satellite altimetry data for studying inland water bodies and river floods. *Journal of Hydrology*, 505, 78–90. <https://doi.org/10.1016/j.jhydrol.2013.09.010>
- Beckley, B., Ray, R., Holmes, S., Zelensky, N., Lemoine, F., Yang, X., et al. (2013). Integrated multi-mission ocean altimeter data for climate research TOPEX/Poseidon, Jason-1, and OSTM/Jason-2 user's handbook, PO. DAAC, CA, USA. *Doi*, 10, 5067.
- Benada, J. (1997). PO. DAAC Merged GDR (TOPEX-POSEIDON) Generation B User's Handbook. Version 2.0. *Jet Propulsion Laboratory, Pasadena, CA*.
- Berry, P. A. M., Garlick, J. D., Freeman, J. A., & Mathers, E. L. (2005). Global inland water monitoring from multi-mission altimetry. *Geophysical Research Letters*, 32(16), n/a-n/a. <https://doi.org/10.1029/2005GL022814>
- Biancamaria, S., Frappart, F., Leleu, A.-S., Marieu, V., Blumstein, D., Desjonquères, J.-D., et al. (2017a). Satellite radar altimetry water elevations performance over a 200 m wide river: Evaluation over

the Garonne River. *Advances in Space Research*, 59(1), 128–146.

<https://doi.org/10.1016/j.asr.2016.10.008>

Biancamaria, S., Frappart, F., Leleu, A.-S., Marieu, V., Blumstein, D., Desjonquères, J.-D., et al. (2017b).

Satellite radar altimetry water elevations performance over a 200m wide river: Evaluation over the Garonne River. *Advances in Space Research*, 59(1), 128–146.

<https://doi.org/10.1016/j.asr.2016.10.008>

Biancamaria, S., Schaedele, T., Blumstein, D., Frappart, F., Boy, F., Desjonquères, J.-D., et al. (2018).

Validation of Jason-3 tracking modes over French rivers. *Remote Sensing of Environment*, 209, 77–89. <https://doi.org/10.1016/j.rse.2018.02.037>

Birkett, C. M., Mertes, L., Dunne, T., Costa, M., & Jasinski, M. (2002). Surface water dynamics in the

Amazon Basin: Application of satellite radar altimetry (DOI 10.1029/2001JD000609). *JOURNAL OF GEOPHYSICAL RESEARCH-ALL SERIES-*, 107(20; SECT 4), LBA-26.

Boergens, E., Dettmering, D., Schwatke, C., & Seitz, F. (2016). Treating the hooking effect in satellite

altimetry data: A case study along the Mekong River and its tributaries. *Remote Sensing*, 8(2), 91.

Carabajal, C. C., Harding, D. J., Boy, J.-P., Danielson, J. J., Gesch, D. B., & Suchdeo, V. P. (2011). Evaluation

of the global multi-resolution terrain elevation data 2010 (GMTED2010) using ICESat geodetic control (pp. 82861Y-82861Y). Presented at the International Symposium on Lidar and Radar Mapping Technologies, International Society for Optics and Photonics.

CASH consortium (IRD, Legos, CLS and BRli, reprocessed and distributed by Legos/CTOH. (2017).

Altimetric data of the TOPEX mission. https://doi.org/10.6096/CTOH_TOPEX_CASH_2017_01

Cole, S., Alan, Buis, & Julien, Watelet. (2013, July 3). Long-Running NASA/CNES Ocean Satellite Takes

Final Bow. Retrieved August 1, 2019, from <https://www.nasa.gov/press/2013/july/long-running-nasacnes-ocean-satellite-takes-final-bow/#.UdREOdhi0gg>

- Constantine, J. A., Dunne, T., Ahmed, J., Legleiter, C., & Lazarus, E. D. (2014). Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. *Nature Geoscience*, 7(12), 899.
- Cooley, S. W., & Pavelsky, T. M. (2016). Spatial and temporal patterns in Arctic river ice breakup revealed by automated ice detection from MODIS imagery. *Remote Sensing of Environment*, 175, 310–322. <https://doi.org/10.1016/j.rse.2016.01.004>
- Danielson, J. J., & Gesch, D. B. (2008). An enhanced global elevation model generalized from multiple higher resolution source datasets. *Int Arch Photogramm Remote Sens Spat Inf Sci*, 37, 1857–1864.
- Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)* (Report No. 2011–1073). Retrieved from <http://pubs.er.usgs.gov/publication/ofr20111073>
- Desjournès, J. D., Carayon, G., Steunou, N., & Lambin, J. (2010). Poseidon-3 Radar Altimeter: New Modes and In-Flight Performances. *Marine Geodesy*, 33(sup1), 53–79. <https://doi.org/10.1080/01490419.2010.488970>
- Duchossois, G. (1991). The ERS-1 mission objectives. *ESA Bulletin*, (65), 16–25.
- Dumont, J., Rosmorduc, V., Picot, N., Desai, S., Bonekamp, H., Figa, J., et al. (2009). OSTM/Jason-2 products handbook. *CNES: SALP-MU-M-OP-15815-CN, EUMETSAT: EUM/OPS-JAS/MAN/08/0041, JPL: OSTM-29-1237, NOAA/NESDIS: Polar Series/OSTM J, 400, 1.*
- Dumont, J., Rosmorduc, V., Carrere, L., Picot, N., Bronner, E., Couhert, A., et al. (2016). *Jason-3 products handbook*. SALP-MU-M-OP-16118-CN.
- Durand, M., Neal, J., Rodríguez, E., Andreadis, K. M., Smith, L. C., & Yoon, Y. (2014). Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope. *Journal of Hydrology*, 511, 92–104.

- Environment Canada. (2016). Surface water data. Inland Waters Directorate, Water Resources Branch, Water Survey of Canada, Ottawa.
- EROS, U. (1996). HYDRO1k Elevation Derivative Database: <http://eros.usgs.gov/#/FindData>. *Products AndDataAvailable/Gtopo30/Hydro* (Last Access: 30 September 2012).
- ERS-2 Product Handbook. (2016), 1.1. https://doi.org/CTOH_ERS-2_2015_01
- Fernandes, J. M., Lázaro, C., Nunes, L. A., & Scharroo, R. (2014). Atmospheric Corrections for Altimetry Studies over Inland Water. *Remote Sensing*, 6(6). <https://doi.org/10.3390/rs6064952>
- Fetterer, F. M., Drinkwater, M. R., Jezek, K. C., Laxon, S. W., Onstott, R. G., & Ulander, L. M. (1992). Sea ice altimetry. *Microwave Remote Sensing of Sea Ice*, 111–135.
- Flohrer, C., Otten, M., Springer, T., & Dow, J. (2011). Generating precise and homogeneous orbits for Jason-1 and Jason-2. *Advances in Space Research*, 48(1), 152–172. <https://doi.org/10.1016/j.asr.2011.02.017>
- Francis, C., Graf, G., Edwards, P., McCraig, M., McCarthy, C., Lefebvre, A., et al. (1995). The ERS-2 spacecraft and its payload. *Esa Bulletin*, 83.
- Frappart, F., Calmant, S., Cauhopé, M., Seyler, F., & Cazenave, A. (2006). Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. *Remote Sensing of Environment*, 100(2), 252–264. <https://doi.org/10.1016/j.rse.2005.10.027>
- Fu, L.-L., & Cazenave, A. (2000). *Satellite altimetry and earth sciences: a handbook of techniques and applications* (Vol. 69). Academic Press.
- Fu, L.-L., Christensen, E. J., Yamarone, C. A., Lefebvre, M., Ménard, Y., Dorrer, M., & Escudier, P. (1994). TOPEX/POSEIDON mission overview. *Journal of Geophysical Research: Oceans*, 99(C12), 24369–24381. <https://doi.org/10.1029/94JC01761>
- Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis*, 54(4), 1167–1178.

- Gleason, C. J., & Smith, L. C. (2014). Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, *111*(13), 4788–4791.
- Gleason, S., Gommenginger, C., & Cromwell, D. (2010). Fading statistics and sensing accuracy of ocean scattered GNSS and altimetry signals. *Advances in Space Research*, *46*(2), 208–220.
- Gommenginger, C., Thibaut, P., Fenoglio-Marc, L., Quartly, G., Deng, X., Gómez-Enri, J., et al. (2011). Retracking Altimeter Waveforms Near the Coasts. In S. Vignudelli, A. G. Kostianoy, P. Cipollini, & J. Benveniste (Eds.), *Coastal Altimetry* (pp. 61–101). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12796-0_4
- Hastings, D. A., & Dunbar, P. (1998). Development & assessment of the Global Land One-km Base Elevation digital elevation model (GLOBE). *Group*, *4*(6), 218–221.
- Hickin, E. J., & Nanson, G. C. (1984). Lateral migration rates of river bends. *Journal of Hydraulic Engineering*, *110*(11), 1557–1567.
- Huggel, C., Schneider, D., Miranda, P. J., Delgado Granados, H., & Kääh, A. (2008). Evaluation of ASTER and SRTM DEM data for lahar modeling: A case study on lahars from Popocatepetl Volcano, Mexico. *The 1994-Present Eruption of Popocatepetl: Background, Current Activity, and Impacts*, *170*(1–2), 99–110. <https://doi.org/10.1016/j.jvolgeores.2007.09.005>
- Hupp, E., & Moreaux, E. (2006). NASA's TOPEX/Poseidon Oceanography Mission Ends. *Press Release*, 06–001.
- Jason-1 Products Handbook. (2016), 5.1. <https://doi.org/SMM-MU-M5-OP-13184-CN>
- Jeffries, M. O., Morris, K., & Kozlenko, N. (2013). Ice Characteristics and Processes, and Remote Sensing of Frozen Rivers and Lakes. In *Remote Sensing in Northern Hydrology: Measuring Environmental Change* (pp. 63–90). American Geophysical Union. <https://doi.org/10.1029/163GM05>

- Kramer, H. J. (2002). *Observation of the Earth and its Environment: Survey of Missions and Sensors*. Springer Science & Business Media.
- Kuo, C.-Y., & Kao, H.-C. (2011). Retracked Jason-2 Altimetry over Small Water Bodies: Case Study of Bajhang River, Taiwan. *Marine Geodesy*, 34(3–4), 382–392.
<https://doi.org/10.1080/01490419.2011.584830>
- Lafon, T., & Parisot, F. (1998). The Jason-1 satellite design and development status.
- Lambin, J., Morrow, R., Fu, L.-L., Willis, J. K., Bonekamp, H., Lillibridge, J., et al. (2010). The OSTM/Jason-2 Mission. *Marine Geodesy*, 33(sup1), 4–25. <https://doi.org/10.1080/01490419.2010.491030>
- Leopold, L. B., & Maddock, T. (1953). *The hydraulic geometry of stream channels and some physiographic implications* (Vol. 252). US Government Printing Office.
- Louet, J., & Bruzzi, S. (1999). ENVISAT mission and system (Vol. 3, pp. 1680–1682). Presented at the Geoscience and Remote Sensing Symposium, 1999. IGARSS'99 Proceedings. IEEE 1999 International, IEEE.
- Maillard, P., Bercher, N., & Calmant, S. (2015). New processing approaches on the retrieval of water levels in Envisat and SARAL radar altimetry over rivers: A case study of the São Francisco River, Brazil. *Remote Sensing of Environment*, 156, 226–241. <https://doi.org/10.1016/j.rse.2014.09.027>
- Michailovsky, C., McEnnis, S., Berry, P., Smith, R., & Bauer-Gottwein, P. (2012). River monitoring from satellite radar altimetry in the Zambezi River basin. *Hydrology and Earth System Sciences*, 16(7), 2181–2192.
- Papa, F., Bala, S. K., Pandey, R. K., Durand, F., Gopalakrishna, V. V., Rahman, A., & Rossow, W. B. (2012). Ganga-Brahmaputra river discharge from Jason-2 radar altimetry: An update to the long-term satellite-derived estimates of continental freshwater forcing flux into the Bay of Bengal. *Journal of Geophysical Research: Oceans*, 117(C11), n/a–n/a. <https://doi.org/10.1029/2012JC008158>

- Pavelsky, T. M., & Smith, L. C. (2004). Spatial and Temporal Patterns in Arctic River Ice Breakup Observed with Modis and Avhrr Time Series. *Remote Sensing of Environment*, 93(3), 328–338.
- Pavelsky, T. M., & Smith, L. C. (2008). RivWidth: A Software Tool for the Calculation of River Widths From Remotely Sensed Imagery. *IEEE Geoscience and Remote Sensing Letters*, 5(1), 70–73.
<https://doi.org/10.1109/LGRS.2007.908305>
- Pavlis, N. K., Holmes, S. A., Kenyon, S. C., & Factor, J. K. (2012). The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *Journal of Geophysical Research: Solid Earth*, 117(B4), n/a–n/a. <https://doi.org/10.1029/2011JB008916>
- Proud, L., & Battrick, B. (1992). ERS-1 User handbook. *ESA SP, Paris: European Space Agency (ESA), 1992*, Edited by Proud, Lois; Battrick, Bruce.
- Rodriguez, E., Morris, C. S., & Belz, J. E. (2006). A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing*, 72(3), 249–260.
- Santos da Silva, J., Calmant, S., Seyler, F., Rotunno Filho, O. C., Cochonneau, G., & Mansur, W. J. (2010). Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote Sensing of Environment*, 114(10), 2160–2181. <https://doi.org/10.1016/j.rse.2010.04.020>
- Schroeder, M., Kornus, W., Lehner, M., Müller, R., Reinartz, P., & Seige, P. (2000). MOMS-the first along track stereo camera in space-a mission review (Vol. 400). Presented at the ASPRS Congress 2000.
- Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI—an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrology and Earth System Sciences*, 19(10), 4345–4364.
- Soussi, B., & Féménias, P. (2009). ENVISAT ALTIMETRY Level 2 User Manual, (1.3). Retrieved from https://earth.esa.int/c/document_library/get_file?folderId=38553&name=DLFE-688.pdf

Sulistioadi, Y. B., Tseng, K. H., Shum, C. K., Hidayat, H., Sumaryono, M., Suhardiman, A., et al. (2015).

Satellite radar altimetry for monitoring small rivers and lakes in Indonesia. *Hydrology and Earth System Sciences*, 19(1), 341–359.

Tourian, M., Tarpanelli, A., Elmi, O., Qin, T., Brocca, L., Moramarco, T., & Sneeuw, N. (2016).

Spatiotemporal densification of river water level time series by multimission satellite altimetry. *Water Resources Research*.

USGS. (2016). USGS Water Data for the Nation.

USGS-Eros data center. (1997). U.S. Geological Survey-EROS Data Center. Retrieved from :

<http://edcdaac.usgs.gov/GTOPO30/README.asp>.

Van Zyl, J. J. (2001). The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. *Acta Astronautica*, 48(5), 559–565.

Wingham, D., Rapley, C., & Griffiths, H. (1986). New techniques in satellite altimeter tracking systems (Vol. 86, pp. 1339–1344). Presented at the Proceedings of IGARSS.

X. Yang, T. M. Pavelsky, G. H. Allen, & G. Donchyts. (2019). RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction From Remotely Sensed Imagery. *IEEE Geoscience and Remote Sensing Letters*, 1–5. <https://doi.org/10.1109/LGRS.2019.2920225>